

Unidad 1. 4 - Gestores de Bases de Datos

- **SXBD = Sistema Gestor de Bases de Datos**
 - Son aplicaciones que permiten a los usuarios **definir, crear y mantener** una o varias bases de datos y proporcionan **acceso controlado** a las mismas.
 - **Mantienen el modelo** de datos de la BD y administran sus datos.
 - Existen SGDB **específicos** para cada modelo de datos (jerárquico, en red, ...), nosotros trabajaremos solo con SGDB relacionales.
 - Para el funcionamiento de los SGDB relacionales se utilizan una serie de lenguajes específicos, incluidos en la definición del modelo relacional y basados en el **álgebra relacional**.
 - Ejemplos de SGDB: PostgreSQL, MySQL, MariaDB, Oracle, ...
- **Servicios que proporciona un SGDB:**
 - **Definición de la base de datos:** Permiten especificar la **estructura** y tipo de los datos de la BD, así como sus restricciones.
 - **Manipulación de los datos:** Consulta, inserción, eliminación y modificación de los datos almacenados en la BD.
 - **Acceso controlado a la BD:** Mediante sistemas que permiten establecer restricciones de acceso y control sobre los datos:
 - Sistema de seguridad que controla el **acceso de los usuarios** a los datos.
 - Sistema de integridad que mantiene la **integridad y consistencia** de los datos.
 - Sistema de **control de concurrencia** que permite el acceso compartido a los datos.
 - **Diccionario de datos** o catálogo de datos que contiene la descripción de los datos de la BD.
- **El modelo de BD incluye la definición de lenguajes específicos para gestionar los datos, que los SGDB utilizan:**
 - **DDL** (Data Definition Language) o Lenguaje de Definición de Datos: Se utilizan para **definir** los datos, sus estructuras, sus tipos y las restricciones entre ellos.
 - **DML** (Data Manipulation Language) o Lenguaje de Manipulación de Datos: Se utiliza para realizar **consultas, inserciones, modificaciones y eliminaciones** sobre los datos de la BD.
 - **DCL** (Data Control Language) o Lenguaje de Control de Datos: Se utiliza para establecer reglas que controlen el **acceso de los usuarios** a

los datos.

- **TCL** (Transaction Control Language) o Lenguaje de Control de Transacciones: Se utiliza para **controlar las transacciones**, que son operaciones que incluyen varios pasos y que deben realizarse todos o ninguno para mantener la integridad de los datos.
- **La arquitectura de BD define:**
 - **Tres niveles de abstracción:** **físico o interno**, **conceptual lógico** , y **externo o de usuario**.
 - **Dos tipos de independencia de datos:** **física** (los cambios en el esquema físico no afectan al esquema lógico) y **lógica** (los cambios en el esquema lógico no afectan a la visión de los usuarios).
- Los SGDB gestionan los tres niveles de abstracción y los dos tipos de independencia.
 - El **modelo físico** (cómo se almacenan los datos en memoria y cómo se accede a ellos) es gestionado internamente por el **SGDB** de forma **transparente** para el usuario.
 - Los **modelos conceptual y de usuario** son definidos y gestionados por el **diseñador y el administrador** de bases de datos.
 - Los **usuarios finales** solo acceden a los datos a través de su propio **modelo externo (vista)**.

Diccionario de datos

- Elemento de la base de datos que recoge la **descripción de todos los elementos de la BD**, sus **tipos de datos** y sus **descripciones**.
- Las descripciones de los datos son a su vez datos almacenados en la BD: **metadatos**.
- Los diccionarios de datos son **accesibles a los usuarios** (según su modelo externo) y se utilizan para obtener **información semántica** sobre los datos.
- **Información almacenada** en el diccionario de datos:
 - **Nombre, tipo y tamaño** de los datos.
 - Nombre de las **relaciones** entre los datos.
 - **Restricciones** de integridad sobre los datos.
 - **Nombres de los usuarios** (personas o aplicaciones) autorizados a acceder a la base de datos.
 - **Esquemas interno, conceptual y externo**, y correspondencia entre ellos.
 - **Estadísticas** de uso.
 - ...
- **Beneficios** que aporta el uso de diccionarios de datos:
 - **Control centralizado** de los datos en un solo punto.
 - Concreción del **significado de los datos**, favoreciendo la comunicación.
 - **Información semántica** de los datos disponible para los usuarios.

- Control de **qué datos pertenecen a cada usuario**.
- Fácil **identificación de redundancias e inconsistencias** de los datos, dado que están centralizados.
- Mantenimiento de un **historial de cambios** de los datos.
- Permite determinar el **impacto de un cambio antes de ser aplicado**.
- Control de la **seguridad**.
- Control de la **integridad** de los datos.

Funciones de los DBMS

- Capacidad de **almacenar datos** en la base de datos, **acceder** a ellos y **actualizarlos**, ocultando al usuario la estructura interna de los datos.
- Proporcionar un **catálogo** en el que se almacenen las descripciones de los datos (**diccionario de datos**) y hacerlo accesible a los usuarios.
- **Control de transaccionalidad**: garantizar que las actualizaciones incluidas en la misma transacción se realicen todas o ninguna.
- **Control de concurrencia**: garantizar que los datos se actualicen correctamente cuando varios usuarios los están actualizando simultáneamente.
- **Capacidad de recuperación** de los datos en caso de que ocurra algún incidente que los dañe (interrupción de operaciones, fallos de hardware, fallos de software, ...).
- **Control de acceso**: solo los usuarios autorizados pueden acceder a los datos.
- Integración con **software de comunicación** que permita el **acceso remoto** a los datos desde otros equipos.
- Permitir establecer **reglas** que garanticen la **integridad y consistencia** de los datos, y que estos se correspondan con su semántica (**restricciones**).
- Mantener la **independencia externa**, de forma que cambios en la estructura de los datos **no afecten a los programas** que los utilizan. Esto es fácil en el caso de añadir elementos o campos, pero difícil en el caso de eliminarlos.
- Proporcionar **herramientas que permitan administrar la base de datos** de forma efectiva. Algunas de estas estarán incluidas por el fabricante, mientras que otras deben ser producidas por el administrador. Por ejemplo, herramientas para: importar y exportar datos, monitorización de uso y funcionamiento, análisis estadístico de uso o de prestaciones, reorganización de índices, gestión del almacenamiento físico, ...

Ventajas de los DBMS

- Ventajas de los Sistemas Gestores de Bases de Datos con respecto a los Sistemas Gestores de Archivos.
 - Debidas a la **integración** de los datos.
 - Debidas a la **interfaz común** que proporcionan.

Ventajas debidas a la integración de los datos

- **Control de redundancia:** evita todos los problemas de la redundancia (costo de almacenamiento, velocidad, pérdida de integridad, ...). Mantienen cierto nivel de redundancia, necesario para relacionar los datos o para mejorar las prestaciones.
- **Consistencia de los datos:** al no haber apenas redundancia y estar el mismo dato accesible para todos los usuarios, es más fácil mantener la consistencia. En el caso de duplicados, es posible establecer restricciones que garanticen la consistencia.
- Posibilidad de obtener más información con los mismos datos, al estar **todos integrados**.
- **Compartición de datos:** los datos no están aislados en los procesos o departamentos que los usan, sino que son compartidos por todos los usuarios, en función de sus permisos.

Ventajas debidas a la interfaz común

- Mejora en la integridad de los datos (validez y consistencia de los datos), mediante el uso de restricciones sobre los datos o las relaciones entre ellos. Algunas restricciones son establecidas por el propio SGDB, mientras que otras deben ser establecidas por el administrador.
- Mejora en la **seguridad** frente a usuarios **no autorizados**, al establecer un control de acceso mediante **claves**, y al establecer un sistema de autorizaciones a nivel de operación y de datos, que permite, por ejemplo, que un usuario pueda consultar un dato pero no modificarlo.
- Mejora de la **accesibilidad de los datos**, mediante el uso de lenguajes de consulta o generación de informes en el **propio SGDB**, sin necesidad de usar un programa adicional.
- **Mejora de la productividad**, al disponer en el propio sistema de muchas **funciones estándar** (suma, unión de datos, subcadenas, ...) que pueden ser utilizadas por el programador sin necesidad de implementarlas.
- Mejora del **mantenimiento** gracias a la **independencia de los datos**: cambios en los datos generalmente no requieren cambiar los programas que los utilizan.
- Aumento de la **conurrencia**, mediante la gestión de la concurrencia, que permite múltiples accesos (consultas, modificaciones, ...) simultáneos a los mismos datos sin que afecte a su integridad.
- Mejora de los servicios de **copias de seguridad y recuperación** ante fallos, mediante su automatización.

Inconvenientes de los DBMS

- **Complejidad:** Es necesario comprender la funcionalidad que ofrecen los SGDB para poder aprovecharlos.
- **Tamaño:** Requieren mucho espacio en disco y mucha memoria para funcionar eficientemente.
- **Costo económico:** Los SGDB suelen ser costosos tanto en términos de licencias como de hardware necesario.

- **Costo del equipo adicional:** Debido a su tamaño y requerimientos, a menudo se necesita hardware adicional para soportar un SGDB.
- **Costo de conversión:** Si se parte de un sistema de gestión de archivos existente, la **migración** a un SGDB puede ser costosa y llevar tiempo.
- Menores prestaciones para algún **proceso específico:** En algunos casos, para ciertas operaciones específicas, los sistemas de gestión de archivos pueden ser más eficientes que los SGDB.
- **Vulnerabilidad a fallos:** Dado que los datos están centralizados en un SGDB, si este falla, toda la organización puede quedar sin acceso a los datos, lo que puede ser un problema crítico.

Lenguajes de los DBMS

El modelo de base de datos incluye la definición de lenguajes específicos para gestionar los datos, que los Sistemas de Gestión de Bases de Datos (SGDB) utilizan:

- **DDL** (Data Definition Language) o Lenguaje de Definición de Datos.
- **DML** (Data Manipulation Language) o Lenguaje de Manipulación de Datos.
- **DCL** (Data Control Language) o Lenguaje de Control de Datos.
- **TCL** (Transaction Control Language) o Lenguaje de Control de Transacciones.

Estos lenguajes permiten proporcionar las funcionalidades necesarias sobre la base de datos para cada tipo de usuario: administradores, diseñadores, programadores de aplicaciones y usuarios finales.

DDL (Data Definition Language) o Lenguaje de Definición de Datos:

- Se utiliza para **definir** los datos, sus estructuras, sus tipos, sus restricciones y las relaciones entre los datos.
- Se utiliza después de diseñar la base de datos y **elegir el Sistema de Gestión de Bases de Datos** (SGDB) que se utilizará, para implementar la base de datos diseñada en el SGDB elegido.
- El lenguaje DDL **afecta a la estructura** de la base de datos, no a los propios datos almacenados.
- Es utilizado por **administradores, diseñadores** del sistema y **programadores**. Nunca es utilizado por los usuarios finales del sistema.
- Se utiliza para **implementar la base de datos** diseñada **en el Sistema de Gestión de Bases de Datos** (SGDB) elegido, lo que implica:
 - **Crear tablas y sus campos**, estableciendo tipos de datos, restricciones, semántica, relaciones con otras tablas, etc.
 - **Crear usuarios** y otorgarles **permisos**.

- **Crear índices** y otras estructuras para mejorar la velocidad.
- **Crear vistas** con la información disponible para cada usuario.
- Y más.
- También se utiliza para realizar **cambios en la estructura** de la base de datos en el **futuro**, lo que incluye:
 - **Añadir o eliminar tablas**, campos, índices, restricciones, etc.
 - **Modificar tipos** de datos, restricciones, permisos, etc.

DML (Data Manipulation Language) o Lenguaje de Manipulación de Datos:

- Se utiliza para realizar **consultas, inserciones, modificaciones y eliminaciones** de **datos** en la base de datos.
- Se utiliza una vez que la **base de datos ya ha sido creada** y su **estructura definida**.
- Es utilizado por **todos los usuarios** de la base de datos, así como por **programas**.
- En bases de datos **jerárquicas y en red**, el lenguaje DML es **procedural**, lo que significa que se debe especificar qué dato se obtiene y qué camino seguir para obtenerlo. A menudo, las sentencias DML deben estar incrustadas en un **lenguaje de alto nivel** para utilizar sus estructuras de control.
- En bases de datos **relacionales**, el lenguaje DML es **declarativo**, lo que significa que el usuario especifica qué dato se obtiene, pero no la forma de acceder a él. El sistema de gestión de bases de datos (**SGDB**) se encarga del acceso a los datos.

DCL (Data Control Language) o Lenguaje de Control de Datos:

- Se utiliza para establecer reglas que controlen el **acceso de los usuarios** a los datos.
- Gestiona los **permisos de acceso y el acceso concurrente** de varios usuarios.

TCL (Transaction Control Language) o Lenguaje de Control de Transacciones:

- Se utiliza para **controlar las transacciones**, que son operaciones que constan de varios pasos y que deben realizarse todos o ninguno para mantener la integridad de los datos.

Clasificación de los DBMS

- **Según el modelo lógico utilizado:** jerárquico, en red, relacional, orientado a objetos, ...

- **Según el número de usuarios concurrentes:** monousuario o multiusuario.
- **Según la distribución física de la BD:** centralizado o distribuido, en función de si toda la BD y el SXBD están en un solo equipo o distribuidos en varios.
- **Según el modelo:** libre (PostgreSQL, MaríaDB, ...) o propietario (Oracle, MySQL, ...).
- **Según el propósito:** de propósito general o específico. Los de propósito específico se utilizan en entornos con requisitos especiales de seguridad o velocidad que justifican el costo de construir un SXBD específico en lugar de usar uno comercial; por ejemplo, se utilizan en el control aéreo.